Krylov Subspace Methods for Large/Sparse Eigenvalue Problems

Tsung-Ming Huang

Department of Mathematics
National Taiwan Normal University, Taiwan

April 17, 2012

Outline

Theorem 1

Let $\mathcal{V}$ be an eigenspace of A and let V be an orthonormal basis for $\mathcal{V}$. Then there is a unique matrix H such that

$$AV = VH.$$

The matrix H is given by

$$H = V^* AV.$$

If (λ, x) is an eigenpair of A with $x \in \mathcal{V}$, then $(\lambda, V^* x)$ is an eigenpair of H. Conversely, if (λ, s) is an eigenpair of H, then (λ, Vs) is an eigenpair of A.

Theorem 2 (Optimal residuals)

Let $[V \ V_{\perp}]$ *be unitary. Let*

$$R = AV - VH \quad \text{and} \quad S^* = V^*A - HV^*.$$

Then $\|R\|$ *and* $\|S\|$ *are minimized when*

$$H = V^*AV,$$

in which case

$$\begin{array}{ll} (a) & \|R\| = \|V_{\perp}^*AV\|, \\ (b) & \|S\| = \|V^*AV_{\perp}\|, \\ (c) & V^*R = 0. \end{array}$$

Definition 3

Let V be orthonormal. Then V^*AV is a Rayleigh quotient of A.

Theorem 4

Let V be orthonormal, A be Hermitian and

$$R = AV - VH.$$

If $\theta_1, \ldots, \theta_k$ are the eigenvalues of H, then there are eigenvalues $\lambda_{j_1}, \ldots, \lambda_{j_k}$ of A such that

$$|\theta_i - \lambda_{j_i}| \leq \|R\|_2 \quad \textit{and} \quad \sqrt{\sum_{i=1}^{k} (\theta_i - \lambda_{j_i})^2} \leq \sqrt{2}\|R\|_F.$$

Suppose the eigenvalue with maximum module is wanted.

Power method

Compute the dominant eigenpair

Disadvantage

At each step it considers only the single vector $A^k u$, which throws away the information contained in the previously generated vectors $u, Au, A^2u, \ldots, A^{k-1}u$.

Definition 5

Let A be of order n and let $u \neq 0$ be an n vector. Then

$$\{u, Au, A^2u, A^3u, \ldots\}$$

is a Krylov sequence based on A and u. We call the matrix

$$K_k(A, u) = \left[\begin{array}{ccccc} u & Au & A^2u & \cdots & A^{k-1}u \end{array}\right]$$

the kth Krylov matrix. The space

$$\mathcal{K}_k(A, u) = \mathcal{R}[K_k(A, u)]$$

is called the kth Krylov subspace.

By the definition of $\mathcal{K}_k(A, u)$, for any vector $v \in \mathcal{K}_k(A, u)$ can be written in the form

$$v = \gamma_1 u + \gamma_2 A u + \cdots + \gamma_k A^{k-1} u \equiv p(A)u,$$

where

$$p(A) = \gamma_1 I + \gamma_2 A + \gamma_3 A^2 + \cdots + \gamma_k A^{k-1}.$$

Assume that $A^\top = A$ and $Ax_i = \lambda_i x_i$ for $i = 1, \ldots, n$. Write u in the form

$$u = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n.$$

Since $p(A)x_i = p(\lambda_i)x_i$, we have

$$p(A)u = \alpha_1 p(\lambda_1)x_1 + \alpha_2 p(\lambda_2)x_2 + \cdots + \alpha_n p(\lambda_n)x_n. \tag{1}$$

If $p(\lambda_i)$ is large compared with $p(\lambda_j)$ for $j \neq i$, then $p(A)u$ is a good approximation to x_i.

Theorem 6

If $x_i^H u \neq 0$ *and* $p(\lambda_i) \neq 0$, *then*

$$\tan \angle(p(A)u, x_i) \leq \max_{j \neq i} \frac{|p(\lambda_j)|}{|p(\lambda_i)|} \tan \angle(u, x_i).$$

Proof: From (1), we have

$$\cos \angle(p(A)u, x_i) = \frac{|x_i^H p(A)u|}{\|p(A)u\|_2 \|x_i\|_2} = \frac{|\alpha_i p(\lambda_i)|}{\sqrt{\sum_{j=1}^n |\alpha_j p(\lambda_j)|^2}}$$

and

$$\sin \angle(p(A)u, x_i) = \frac{\sqrt{\sum_{j \neq i} |\alpha_j p(\lambda_j)|^2}}{\sqrt{\sum_{j=1}^n |\alpha_j p(\lambda_j)|^2}}$$

Hence

$$\begin{aligned}\tan^2\angle(p(A)u,x_i) &= \sum_{j\neq i}\frac{|\alpha_j p(\lambda_j)|^2}{|\alpha_i p(\lambda_i)|^2}\\ &\leq \max_{j\neq i}\frac{|p(\lambda_j)|^2}{|p(\lambda_i)|^2}\sum_{j\neq i}\frac{|\alpha_j|^2}{|\alpha_i|^2}\\ &= \max_{j\neq i}\frac{|p(\lambda_j)|^2}{|p(\lambda_i)|^2}\tan^2\angle(u,x_i).\end{aligned}$$

■

Assume that $p(\lambda_i)=1$, then

$$\tan\angle(p(A)u,x_i)\leq \max_{j\neq i,p(\lambda_i)=1}|p(\lambda_j)|\tan\angle(u,x_i)\quad\forall\quad p(A)u\in\mathcal{K}_k.$$

Hence

$$\tan\angle(x_i,\mathcal{K}_k)\leq \min_{deg(p)\leq k-1,p(\lambda_i)=1}\max_{j\neq i}|p(\lambda_j)|\tan\angle(u,x_i).$$

Assume that

$$\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$$

and that our interest is in the eigenvector x_1. Then

$$\tan \angle(x_1, \mathcal{K}_k) \leq \min_{deg(p) \leq k-1, p(\lambda_1)=1} \max_{\lambda \in [\lambda_n, \lambda_2]} |p(\lambda)| \tan \angle(u, x_1).$$

Question

How to compute

$$\min_{deg(p) \leq k-1, p(\lambda_1)=1} \max_{\lambda \in [\lambda_n, \lambda_2]} |p(\lambda)|?$$

Definition 7

The Chebyshev polynomials are defined by

$$c_k(t) = \begin{cases} \cos(k \cos^{-1} t), & |t| \leq 1, \\ \cosh(k \cosh^{-1} t), & |t| \geq 1. \end{cases}$$

Theorem 8

(i) $c_0(t) = 1,\ c_1(t) = t$ *and*

$$c_{k+1}(t) = 2c_k(t) - c_{k-1}(t), \quad k = 1, 2, \ldots.$$

(ii) *For* $|t| > 1$, $c_k(t) = (1 + \sqrt{t^2 - 1})^k + (1 + \sqrt{t^2 - 1})^{-k}$.

(iii) *For* $t \in [-1, 1], \quad |c_k(t)| \leq 1$. *Moreover, if*

$$t_{ik} = \cos \frac{(k-i)\pi}{k}, \quad i = 0, 1, \ldots, k,$$

then $\quad c_k(t_{ik}) = (-1)^{k-i}$.

(iv) *For* $s > 1$,

$$\min_{deg(p) \leq k, p(s)=1} \max_{t \in [0,1]} |p(t)| = \frac{1}{c_k(s)}, \tag{2}$$

and the minimum is obtained only for $p(t) = c_k(t)/c_k(s)$.

For applying (2), we define

$$\lambda = \lambda_2 + (\mu - 1)(\lambda_2 - \lambda_n)$$

to transform interval $[\lambda_n, \lambda_2]$ to $[0, 1]$. Then the value of μ at λ_1 is

$$\mu_1 = 1 + \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}$$

and

$$\begin{aligned} & \min_{deg(p) \leq k-1, p(\lambda_1)=1} \max_{\lambda \in [\lambda_n, \lambda_2]} |p(\lambda)| \\ = & \min_{deg(p) \leq k-1, p(\mu_1)=1} \max_{\mu \in [0,1]} |p(\mu)| = \frac{1}{c_{k-1}(\mu_1)} \end{aligned}$$

Theorem 9

Let $A^\top = A$ *and* $Ax_i = \lambda_i x_i,\ i = 1, \cdots, n$ *with* $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n$. *Let* $\eta = \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}$. *Then*

$$\begin{aligned} \tan\angle[x_1, \mathcal{K}_k(A,u)] &\leq \frac{\tan\angle(x_1,u)}{c_{k-1}(1+\eta)} \\ &= \frac{\tan\angle(x_1,u)}{(1+\sqrt{2\eta+\eta^2})^{k-1} + (1+\sqrt{2\eta+\eta^2})^{1-k}}. \end{aligned}$$

- For k large and if η is small, then the bound becomes

$$\tan \angle[x_1, \mathcal{K}_k(A, u)] \lesssim \frac{\tan \angle(x_1, u)}{(1+\sqrt{2\eta})^{k-1}}.$$

- Compare it with power method: If $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, then the conv. rate is $|\lambda_2/\lambda_1|^k$.
- For example, let $\lambda_1 = 1$, $\lambda_2 = 0.95$, $\lambda_3 = 0.95^2$, $\cdots$, $\lambda_{100} = 0.95^{99}$ be the Ews of $A \in \mathbb{R}^{100\times 100}$. Then $\eta = 0.0530$ and the bound on the conv. rate is $1/(1+\sqrt{2\eta}) = 0.7544$. Thus the square root effect gives a great improvement over the rate of 0.95 for the power method.

Definition 10

A Householder transformation or elementary reflector is a matrix of

$$H = I - uu^*$$

where $\|u\|_2 = \sqrt{2}$.

Note that H is Hermitian and unitary.

Theorem 11

Let x be a vector with $x_1 \neq 0$. Let

$$u = \frac{\rho \frac{x}{\|x\|_2} + e_1}{\sqrt{1 + \rho \frac{x_1}{\|x\|_2}}},$$

where $\rho = \bar{x}_1 / |x_1|$. Then

$$Hx = -\bar{\rho} \|x\|_2 e_1.$$

Proof: Since

$$\begin{aligned} & [\bar{\rho} x^*/\|x\|_2 + e_1^\top][\rho x/\|x\|_2 + e_1] \\ = \; & \bar{\rho}\rho + \rho x_1/\|x\|_2 + \bar{\rho}\bar{x}_1/\|x\|_2 + 1 \\ = \; & 2[1 + \rho x_1/\|x\|_2], \end{aligned}$$

it follows that

$$u^* u = 2 \quad \Rightarrow \quad \|u\|_2 = \sqrt{2}$$

and

$$u^* x = \frac{\bar{\rho}\|x\|_2 + x_1}{\sqrt{1 + \rho \frac{x_1}{\|x\|_2}}}.$$

Hence,

$$
\begin{aligned}
Hx &= x - (u^* x) u = x - \frac{\bar{\rho}\|x\|_2 + x_1}{\sqrt{1 + \rho \frac{x_1}{\|x\|_2}}} \frac{\rho \frac{x}{\|x\|_2} + e_1}{\sqrt{1 + \rho \frac{x_1}{\|x\|_2}}} \\
&= \left[1 - \frac{(\bar{\rho}\|x\|_2 + x_1) \frac{\rho}{\|x\|_2}}{1 + \rho \frac{x_1}{\|x\|_2}} \right] x - \frac{\bar{\rho}\|x\|_2 + x_1}{1 + \rho \frac{x_1}{\|x\|_2}} e_1 \\
&= -\frac{\bar{\rho}\|x\|_2 + x_1}{1 + \rho \frac{x_1}{\|x\|_2}} e_1 \\
&= -\bar{\rho}\|x\|_2 e_1.
\end{aligned}
$$

■

Definition 12

A complex $m \times n$-matrix $R = [r_{ij}]$ is called an upper (lower) triangular matrix, if $r_{ij} = 0$ for $i > j$ $(i < j)$.

Definition 13

Given $A \in \mathbb{C}^{m \times n}, Q \in \mathbb{C}^{m \times m}$ unitary and $R \in \mathbb{C}^{m \times n}$ upper triangular such that $A = QR$. Then the product is called a QR-factorization of A.

Theorem 14

Any complex $m \times n$ matrix A can be factorized by the product $A = QR$, where Q is $m \times m$-unitary and R is $m \times n$ upper triangular.

Proof: Let $A^{(0)} = A = [a_1^{(0)} | a_2^{(0)} | \cdots | a_n^{(0)}]$. Find $Q_1 = (I - 2w_1 w_1^*)$ such that $Q_1 a_1^{(0)} = ce_1$. Then

$$\begin{aligned} A^{(1)} &= Q_1 A^{(0)} = [Q_1 a_1^{(0)}, Q_1 a_2^{(0)}, \cdots, Q_1 a_n^{(0)}] \\ &= \left[\begin{array}{c|c|c|c} c_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & a_2^{(1)} & \cdots & a_n^{(1)} \\ 0 & & & \end{array}\right]. \end{aligned} \tag{3}$$

Find $Q_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & I - w_2 w_2^* \end{array}\right]$ such that $(I - 2w_2 w_2^*) a_2^{(1)} = c_2 e_1$. Then

$$A^{(2)} = Q_2 A^{(1)} = \left[\begin{array}{cc|ccc} c_1 & * & * & \cdots & * \\ 0 & c_2 & * & \cdots & * \\ \hline 0 & 0 & & & \\ \vdots & \vdots & a_3^{(2)} & \cdots & a_n^{(2)} \\ 0 & 0 & & & \end{array}\right].$$

We continue this process. Then after $l = \min(m, n)$ steps $A^{(l)}$ is an upper triangular matrix satisfying

$$A^{(l-1)} = R = Q_{l-1} \cdots Q_1 A.$$

Then $A = QR$, where $Q = Q_1^* \cdots Q_{l-1}^*$. ■

Suppose that the columns of K_{k+1} are linearly independent and let

$$K_{k+1} = U_{k+1} R_{k+1}$$

be the QR factorization of K_{k+1}. Then the columns of U_{k+1} are results of successively orthogonalizing the columns of K_{k+1}.

Theorem 15

Let $\|u_1\|_2 = 1$ *and the columns of* $K_{k+1}(A, u_1)$ *be linearly indep. Let* $U_{k+1} = [u_1 \ \cdots \ u_{k+1}]$ *be the* Q*-factor of* K_{k+1}*. Then there is a* $(k+1) \times k$ *unreduced upper Hessenberg matrix*

$$\hat{H}_k \equiv \left[\begin{array}{cccc} \hat{h}_{11} & \cdots & \cdots & \hat{h}_{1k} \\ \hat{h}_{21} & \hat{h}_{22} & \cdots & \hat{h}_{2k} \\ & \ddots & \ddots & \vdots \\ & & \hat{h}_{k,k-1} & \hat{h}_{kk} \\ \hline & & & \hat{h}_{k+1,k} \end{array}\right] \quad \text{with} \quad \hat{h}_{i+1,i} \neq 0 \qquad (4)$$

such that

$$AU_k = U_{k+1}\hat{H}_k. \text{ (Arnoldi decomp.)} \qquad (5)$$

Conversely, if U_{k+1} *is orthonormal and satisfies (5), where* $\hat{H}_k$ *is defined in (4), then* U_{k+1} *is the* Q*-factor of* $K_{k+1}(A, u_1)$*.*

Proof: ("$\Rightarrow$") Let $K_k = U_k R_k$ be the QR factorization and $S_k = R_k^{-1}$. Then

$$AU_k = AK_k S_k = K_{k+1} \begin{bmatrix} 0 \\ S_k \end{bmatrix} = U_{k+1} R_{k+1} \begin{bmatrix} 0 \\ S_k \end{bmatrix} = U_{k+1} \hat{H}_k,$$

where

$$\hat{H}_k = R_{k+1} \begin{bmatrix} 0 \\ S_k \end{bmatrix}.$$

It implies that $\hat{H}_k$ is a $(k+1) \times k$ Hessenberg matrix and

$$h_{i+1,i} = r_{i+1,i+1} s_{ii} = \frac{r_{i+1,i+1}}{r_{ii}}.$$

Thus by the nonsingularity of R_k, $\hat{H}_k$ is unreduced.

("$\Leftarrow$") If $k = 1$, then

$$Au_1 = h_{11}u_1 + h_{21}u_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix}$$

which implies that

$$K_2 = \begin{bmatrix} u_1 & Au_1 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & h_{11} \\ 0 & h_{21} \end{bmatrix}.$$

Since $\begin{bmatrix} u_1 & u_2 \end{bmatrix}$ is orthonormal, $\begin{bmatrix} u_1 & u_2 \end{bmatrix}$ is the Q-factor of K_2. Assume U_k is the Q-factor of K_k, i.e., $K_k = U_k R_k$. By the definition of the Krylov matrix, we have

$$\begin{aligned} K_{k+1} &= \begin{bmatrix} u_1 & AK_k \end{bmatrix} = \begin{bmatrix} u_1 & AU_kR_k \end{bmatrix} = \begin{bmatrix} u_1 & U_{k+1}\hat{H}_kR_k \end{bmatrix} \\ &= U_{k+1}\begin{bmatrix} e_1 & \hat{H}_kR_k \end{bmatrix} \end{aligned}$$

Hence U_{k+1} is the Q-factor of K_{k+1}. $\blacksquare$

- The uniqueness of Hessenberg reduction

Definition 16

Let H be upper Hessenberg of order n. Then H is unreduced if $h_{i+1,i} \neq 0$ for $i = 1, \cdots, n-1$.

Theorem 17 (Implicit Q theorem)

Suppose $Q = \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix}$ *and* $V = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ *are unitary matrices with*

$$Q^* A Q = H \quad \text{and} \quad V^* A V = G$$

being upper Hessenberg. Let k *denote the smallest positive integer for which* $h_{k+1,k} = 0$, *with the convection that* $k = n$ *if* H *is unreduced. If* $v_1 = q_1$, *then* $v_i = \pm q_i$ *and* $|h_{i,i-1}| = |g_{i,i-1}|$ *for* $i = 2, \cdots, k$. *Moreover, if* $k < n$, *then* $g_{k+1,k} = 0$.

Definition 18

Let $U_{k+1} \in \mathbb{C}^{n\times(k+1)}$ be orthonormal. If there is a $(k+1)\times k$ unreduced upper Hessenberg matrix $\hat{H}_k$ such that

$$AU_k = U_{k+1}\hat{H}_k, \tag{6}$$

then (6) is called an Arnoldi decomposition of order k. If $\hat{H}_k$ is reduced, we say the Arnoldi decomposition is reduced.

Partition

$$\hat{H}_k = \begin{bmatrix} H_k \\ h_{k+1,k}e_k^T \end{bmatrix},$$

and set

$$\beta_k = h_{k+1,k}.$$

Then (6) is equivalent to

$$AU_k = U_k H_k + \beta_k u_{k+1} e_k^T.$$

Theorem 19

Suppose the Krylov sequence $K_{k+1}(A, u_1)$ does not terminate at $k+1$. Then up to scaling of the columns of U_{k+1}, the Arnoldi decomposition of K_{k+1} is unique.

Proof: Since the Krylov sequence $K_{k+1}(A, u_1)$ does not terminate at $k+1$, the columns of $K_{k+1}(A, u_1)$ are linearly independent. By Theorem 15, there is an unreduced matrix H_k and $\beta_k \neq 0$ such that

$$AU_k = U_k H_k + \beta_k u_{k+1} e_k^T, \tag{7}$$

where $U_{k+1} = [U_k \ u_{k+1}]$ is an orthonormal basis for $\mathcal{K}_{k+1}(A, u_1)$. Suppose there is another orthonormal basis $\tilde{U}_{k+1} = [\tilde{U}_k \ \tilde{u}_{k+1}]$ for $\mathcal{K}_{k+1}(A, u_1)$, unreduced matrix $\tilde{H}_k$ and $\tilde{\beta}_k \neq 0$ such that

$$A\tilde{U}_k = \tilde{U}_k \tilde{H}_k + \tilde{\beta}_k \tilde{u}_{k+1} e_k^T.$$

Then we claim that

$$\tilde{U}_k^H u_{k+1} = 0.$$

For otherwise there is a column $\tilde{u}_j$ of $\tilde{U}_k$ such that

$$\tilde{u}_j = \alpha u_{k+1} + U_k a, \quad \alpha \neq 0.$$

Hence

$$A\tilde{u}_j = \alpha A u_{k+1} + A U_k a$$

which implies that $A\tilde{u}_j$ contains a component along $A^{k+1}u_1$. Since the Krylov sequence $K_{k+1}(A, u_1)$ does not terminate at $k+1$, we have

$$\mathcal{K}_{k+2}(A, u_1) \neq \mathcal{K}_{k+1}(A, u_1).$$

Therefore, $A\tilde{u}_j$ lies in $\mathcal{K}_{k+2}(A, u_1)$ but not in $\mathcal{K}_{k+1}(A, u_1)$ which is a contradiction.

Since U_{k+1} and $\tilde{U}_{k+1}$ are orthonormal bases for $\mathcal{K}_{k+1}(A, u_1)$ and $\tilde{U}_k^H u_{k+1} = 0$, it follows that

$$\mathcal{R}(U_k) = \mathcal{R}(\tilde{U}_k) \quad \text{and} \quad U_k^H \tilde{u}_{k+1} = 0,$$

that is

$$U_k = \tilde{U}_k Q$$

for some unitary matrix Q. Hence

$$A(\tilde{U}_k Q) = (\tilde{U}_k Q)(Q^H \tilde{H}_k Q) + \tilde{\beta}_k \tilde{u}_{k+1}(e_k^T Q),$$

or

$$AU_k = U_k(Q^H \tilde{H}_k Q) + \tilde{\beta}_k \tilde{u}_{k+1} e_k^T Q. \tag{8}$$

On premultiplying (7) and (8) by U_k^H, we obtain

$$H_k = U_k^H A U_k = Q^H \tilde{H}_k Q.$$

Similarly, premultiplying by u_{k+1}^H, we obtain

$$\beta_k e_k^T = u_{k+1}^H A U_k = \tilde{\beta}_k (u_{k+1}^H \tilde{u}_{k+1}) e_k^T Q.$$

It follows that the last row of Q is $\omega_k e_k^T$, where $|\omega_k| = 1$. Since the norm of the last column of Q is one, the last column of Q is $\omega_k e_k$. Since H_k is unreduced, it follows from the implicit Q theorem that

$$Q = \text{diag}(\omega_1, \cdots, \omega_k), \quad |\omega_j| = 1, \ j = 1, \ldots, k.$$

Thus up to column scaling $U_k = \tilde{U}_k Q$ is the same as $\tilde{U}_k$. Subtracting (8) from (7), we find that

$$\beta_k u_{k+1} = \omega_k \tilde{\beta}_k \tilde{u}_{k+1}$$

so that up to scaling u_{k+1} and $\tilde{u}_{k+1}$ are the same. ■

Let A be Hermitian and let

$$AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top \tag{9}$$

be an Arnoldi decomposition. Since T_k is upper Hessenberg and $T_k = U_k^H A U_k$ is Hermitian, it follows that T_k is tridiagonal and can be written in the form

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{k-2} & \alpha_{k-1} & \beta_{k-1} \\ & & & & \beta_{k-1} & \alpha_k \end{bmatrix}.$$

Equation (9) is called a Lanczos decomposition.

The first column of (9) is

$$Au_1 = \alpha_1 u_1 + \beta_1 u_2 \quad \text{or} \quad u_2 = \frac{Au_1 - \alpha_1 u_1}{\beta_1}.$$

From the orthonormality of u_1 and u_2, it follows that

$$\alpha_1 = u_1^* A u_1$$

and

$$\beta_1 = \|Au_1 - \alpha_1 u_1\|_2.$$

More generality, from the j-th column of (9) we get the relation

$$u_{j+1} = \frac{Au_j - \alpha_j u_j - \beta_{j-1} u_{j-1}}{\beta_j}$$

where

$$\alpha_j = u_j^* A u_j \quad \text{and} \quad \beta_j = \|Au_j - \alpha_j u_j - \beta_{j-1} u_{j-1}\|_2.$$

This is the Lanczos three-term recurrence.

Algorithm 1 (Lanczos recurrence)

Let u_1 be given. This algorithm generates the Lanczos decomposition

$$AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top$$

where T_k is symmetric tridiagonal.

1. $u_0 = 0; \beta_0 = 0;$
2. *for* $j = 1$ *to* k
3. $\quad u_{j+1} = Au_j$
4. $\quad \alpha_j = u_j^* u_{j+1}$
5. $\quad v = u_{j+1} - \alpha_j u_j - \beta_{j-1} u_{j-1}$
6. $\quad \beta_j = \|v\|_2$
7. $\quad u_{j+1} = v/\beta_j$
8. *end for* j

Reorthogonalization

Let

$$\tilde{u}_{j+1} = Au_j - \alpha_j u_j - \beta_{j-1} u_{j-1}$$

with $\alpha_j = u_j^* A u_j$. Re-orthogonalize $\tilde{u}_{j+1}$ against U_j, i.e.,

$$\begin{aligned}\tilde{u}_{j+1} &:= \tilde{u}_{j+1} - \sum_{i=1}^{j} (u_i^* \tilde{u}_{j+1})\, u_i \\ &= Au_j - \left(\alpha_j + u_j^* \tilde{u}_{j+1}\right) u_j - \beta_{j-1} u_{j-1} - \sum_{i=1}^{j-1} (u_i^* \tilde{u}_{j+1})\, u_i\end{aligned}$$

Take

$$\beta_j = \|\tilde{u}_{j+1}\|_2, \quad u_{j+1} = \tilde{u}_{j+1}/\beta_j.$$

Theorem 20 (Stop criterion)

Suppose that j steps of the Lanczos algorithm have been performed and that

$$S_j^H T_j S_j = diag(\theta_1, \cdots, \theta_j)$$

is the Schur decomposition of the tridiagonal matrix T_j, if $Y_j \in \mathbb{C}^{n \times j}$ is defined by

$$Y_j \equiv \begin{bmatrix} y_1 & \cdots & y_j \end{bmatrix} = U_j S_j$$

then for $i = 1, \cdots, j$ we have

$$\|Ay_i - \theta_i y_i\|_2 = |\beta_j||s_{ji}|$$

where $S_j = [s_{pq}]$.

Proof : Post-multiplying

$$AU_j = U_j T_j + \beta_j u_{j+1} e_j^\top$$

by S_j gives

$$AY_j = Y_j diag(\theta_1, \cdots, \theta_j) + \beta_j u_{j+1} e_j^\top S_j,$$

i.e.,

$$Ay_i = \theta_i y_i + \beta_j u_{j+1} (e_j^\top S_j e_i) \,, \quad i = 1, \cdots, j.$$

The proof is complete by taking norms. ■

Remark 1

- *Stop criterion =* $|\beta_j||s_{ji}|$. *Do not need to compute* $\|Ay_i - \theta_i y_i\|_2$.
- *In general,* $|\beta_j|$ *is not small. It is possible that* $|\beta_j||s_{ji}|$ *is small.*

Theorem 21

Let A be $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ and corresponding orthonormal eigenvectors $z_1, \cdots, z_n$. If $\theta_1 \geq \cdots \geq \theta_j$ are the eigenvalues of T_j obtained after j steps of the Lanczos iteration, then

$$\lambda_1 \geq \theta_1 \geq \lambda_1 - \frac{(\lambda_1 - \lambda_n)(\tan\phi_1)^2}{[c_{j-1}(1+2\rho_1)]^2},$$

where $\cos\phi_1 = |u_1^\top z_1|$, c_{j-1} is a Chebychev polynomal of degree $j-1$ and

$$\rho_1 = \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}.$$

Proof: From Courant-Fischer theorem we have

$$\theta_1 = \max_{y\neq 0} \frac{y^\top T_j y}{y^\top y} = \max_{y\neq 0} \frac{(U_j y)^\top A (U_j y)}{(U_j y)^\top (U_j y)} = \max_{0\neq w\in \mathcal{K}_j(u_1,A)} \frac{w^\top A w}{w^\top w}.$$

Since λ_1 is the maximum of $w^\top Aw/w^\top w$ over all nonzero w, it follows that $\lambda_1 \geq \theta_1$. To obtain the lower bound for θ_1, note that

$$\theta_1 = \max_{p\in P_{j-1}} \frac{u_1^\top p(A) A p(A) u_1}{u_1^\top p(A)^2 u_1},$$

where P_{j-1} is the set of all $j-1$ degree polynomials. If $u_1 = \sum_{i=1}^n d_i z_i$, then

$$\begin{aligned} \frac{u_1^\top p(A) A p(A) u_1}{u_1^\top p(A)^2 u_1} &= \frac{\sum_{i=1}^n d_i^2 p(\lambda_i)^2 \lambda_i}{\sum_{i=1}^n d_i^2 p(\lambda_i)^2} \\ &\geq \lambda_1 - (\lambda_1 - \lambda_n) \frac{\sum_{i=2}^n d_i^2 p(\lambda_i)^2}{d_1^2 p(\lambda_1)^2 + \sum_{i=2}^n d_i^2 p(\lambda_i)^2}. \end{aligned}$$

We can make the lower bound tight by selecting a polynomial $p(\alpha)$ that is large at $\alpha = \lambda_1$ in comparison to its value at the remaining eigenvalues. Set

$$p(\alpha) = c_{j-1}\left(-1 + 2\frac{\alpha - \lambda_n}{\lambda_2 - \lambda_n}\right),$$

where $c_{j-1}(z)$ is the $(j-1)$-th Chebychev polynomial generated by

$$c_j(z) = 2zc_{j-1}(z) - c_{j-2}(z), \quad c_0 = 1, c_1 = z.$$

These polynomials are bounded by unity on [-1,1]. It follows that $|p(\lambda_i)|$ is bounded by unity for $i = 2, \cdots, n$ while $p(\lambda_1) = c_{j-1}(1 + 2\rho_1)$. Thus,

$$\theta_1 \geq \lambda_1 - (\lambda_1 - \lambda_n)\frac{(1 - d_1^2)}{d_1^2}\frac{1}{c_{j-1}^2(1 + 2\rho_1)}.$$

The desired lower bound is obtained by noting that $\tan(\phi_1)^2 = (1 - d_1^2)/d_1^2$.

Theorem 22

Using the same notation as Theorem 21,

$$\lambda_n \leq \theta_j \leq \lambda_n + \frac{(\lambda_1 - \lambda_n)\tan^2\varphi_n}{[c_{j-1}(1+2\rho_n)]^2},$$

where

$$\rho_n = \frac{\lambda_{n-1} - \lambda_n}{\lambda_1 - \lambda_{n-1}}, \quad \cos\varphi_n = |u_1^\top z_n|.$$

Proof : Apply Theorem 21 with A replaced by $-A$. ∎

Restarted Lanczos method

Let

$$AU_m = U_m T_m + \beta_m u_{m+1} e_m^T$$

be a Lanczos decomposition.

1. In principle, we can keep expanding the Lanczos decomposition until the Ritz pairs have converged.
2. Unfortunately, it is limited by the amount of memory to storage of U_m.
3. Restarted the Lanczos process once m becomes so large that we cannot store U_m.
   - Implicitly restarting method
   - Krylov-Schur decomposition

Implicitly restarted Lanczos method

- Choose a new starting vector for the underlying Krylov sequence
- A natural choice would be a linear combination of Ritz vectors that we are interested in.

Filter polynomials

Assume A has a complete system of eigenpairs (λ_i, x_i) and we are interested in the first k of these eigenpairs. Expand u_1 in the form

$$u_1 = \sum_{i=1}^{k} \gamma_i x_i + \sum_{i=k+1}^{n} \gamma_i x_i.$$

If p is any polynomial, we have

$$p(A)u_1 = \sum_{i=1}^{k} \gamma_i p(\lambda_i) x_i + \sum_{i=k+1}^{n} \gamma_i p(\lambda_i) x_i.$$

- Choose p so that the values $p(\lambda_i)$ $(i = k+1, \ldots, n)$ are small compared to the values $p(\lambda_i)$ $(i = 1, \ldots, k)$.
- Then $p(A)u_1$ is rich in the components of the x_i that we want and deficient in the ones that we do not want.
- p is called a filter polynomial.
- Suppose we have Ritz values $\theta_1, \ldots, \theta_m$ and $\theta_1, \ldots, \theta_{m-k}$ are not interesting. Then take

$$p(t) = (t - \theta_1) \cdots (t - \theta_{m-k}).$$

Implicitly restarted Lanczos: Let

$$AU_m = U_m T_m + \beta_m u_{m+1} e_m^\top \tag{10}$$

be a Lanczos decomposition with order m. Choose a filter polynomial p of degree $m-k$ and use the implicit restarting process to reduce the decomposition to a decomposition

$$A\tilde{U}_k = \tilde{U}_k \tilde{T}_k + \tilde{\beta}_k \tilde{u}_{k+1} e_k^\top$$

of order k with starting vector $p(A)u_1$.

Let $\theta_1, \ldots, \theta_m$ be eigenvalues of T_m and suppose that $\theta_1, \ldots, \theta_{m-k}$ correspond to the part of the spectrum we are not interested in. Then take

$$p(t) = (t - \theta_1)(t - \theta_2) \cdots (t - \theta_{m-k}).$$

The starting vector $p(A)u_1$ is equal to

$$\begin{aligned} p(A)u_1 &= (A - \theta_{m-k}I) \cdots (A - \theta_2 I)(A - \theta_1 I)u_1 \\ &= (A - \theta_{m-k}I)\left[\cdots\left[(A - \theta_2 I)\left[(A - \theta_1 I)u_1\right]\right]\right]. \end{aligned}$$

In the first, we construct a Lanczos decomposition with starting vector $(A - \theta_1 I)u_1$. From (10), we have

$$\begin{aligned} (A - \theta_1 I)U_m &= U_m(T_m - \theta_1 I) + \beta_m u_{m+1} e_m^\top \\ &= U_m Q_1 R_1 + \beta_m u_{m+1} e_m^\top, \end{aligned} \tag{11}$$

where

$$T_m - \theta_1 I = Q_1 R_1$$

is the QR factorization of $T_m - \theta_1 I$.

Postmultiplying by Q_1, we get

$$(A - \theta_1 I)(U_m Q_1) = (U_m Q_1)(R_1 Q_1) + \beta_m u_{m+1}(e_m^\top Q_1).$$

It implies that

$$AU_m^{(1)} = U_m^{(1)} T_m^{(1)} + \beta_m u_{m+1} b_{m+1}^{(1)H},$$

where

$$U_m^{(1)} = U_m Q_1, \quad T_m^{(1)} = R_1 Q_1 + \theta_1 I, \quad b_{m+1}^{(1)H} = e_m^\top Q_1.$$

($T_m^{(1)}$: one step of single shifted QR algorithm)

Theorem 23

Let T_m *be an unreduced tridiagonal. Then* $T_m^{(1)}$ *has the form*

$$T_m^{(1)} = \begin{bmatrix} \hat{T}_m^{(1)} & 0 \\ 0 & \theta_1 \end{bmatrix},$$

where $\hat{T}_m^{(1)}$ *is unreduced tridiagonal.*

Proof: Let

$$T_m - \theta_1 I = Q_1 R_1$$

be the QR factorization of $T_m - \theta_1 I$ with

$$Q_1 = G(1, 2, \varphi_1) \cdots G(m-1, m, \varphi_{m-1})$$

where $G(i, i+1, \varphi_i)$ for $i = 1, \ldots, m-1$ are Given rotations.

Since T_m is unreduced tridiagonal, i.e., the subdiagonal elements of T_m are nonzero, we get

$$\varphi_i \neq 0 \quad \text{for } i = 1, \ldots, m-1 \tag{12}$$

and

$$(R_1)_{ii} \neq 0 \quad \text{for } i = 1, \ldots, m-1. \tag{13}$$

Since θ_1 is an eigenvalue of T_m, we have that $T_m - \theta_1 I$ is singular and then

$$(R_1)_{mm} = 0. \tag{14}$$

Using the results of (12), (13) and (14), we get

$$\begin{aligned} T_m^{(1)} &= R_1 Q_1 + \theta_1 I = R_1 G(1, 2, \varphi_1) \cdots G(m-1, m, \varphi_{m-1}) + \theta_1 I \\ &= \begin{bmatrix} \hat{H}_m^{(1)} & \hat{h}_{12} \\ 0 & \theta_1 \end{bmatrix}, \end{aligned} \tag{15}$$

where $\hat{H}_m^{(1)}$ is unreduced upper Hessenberg.

By the definition of $T_m^{(1)}$, we get

$$Q_1 T_m^{(1)} Q_1^H = Q_1(R_1 Q_1 + \theta_1 I) Q_1^H = Q_1 R_1 + \theta_1 I = T_m.$$

It implies that $T_m^{(1)}$ is tridiagonal and then, from (15), the result in (12) is obtained. ■

Remark 2

- *$U_m^{(1)}$ is orthonormal.*
- *The vector $b_{m+1}^{(1)H} = e_m^\top Q_1$ has the form*

$$b_{m+1}^{(1)H} = \begin{bmatrix} 0 & \cdots & 0 & q_{m-1,m}^{(1)} & q_{m,m}^{(1)} \end{bmatrix};$$

i.e., only the last two components of $b_{m+1}^{(1)}$ are nonzero.

- For on postmultiplying (11) by e_1, we get

$$(A - \theta_1 I)u_1 = (A - \theta_1 I)(U_m e_1) = U_m^{(1)} R_1 e_1 = r_{11}^{(1)} u_1^{(1)}.$$

Since T_m is unreduced, $r_{11}^{(1)}$ is nonzero. Therefore, the first column of $U_m^{(1)}$ is a multiple of $(A - \theta_1 I)u_1$.

- By the definition of $T_m^{(1)}$, we get

$$Q_1 T_m^{(1)} Q_1^H = Q_1(R_1 Q_1 + \theta_1 I)Q_1^H = Q_1 R_1 + \theta_1 I = T_m.$$

Therefore, $\theta_1, \theta_2, \ldots, \theta_m$ are also eigenvalues of $T_m^{(1)}$.

Similarly,

$$\begin{aligned}(A-\theta_2 I)U_m^{(1)} &= U_m^{(1)}(T_m^{(1)}-\theta_2 I)+\beta_m u_{m+1}b_{m+1}^{(1)H} \\ &= U_m^{(1)}Q_2R_2+\beta_m u_{m+1}b_{m+1}^{(1)H},\end{aligned} \tag{16}$$

where

$$T_m^{(1)}-\theta_2 I = Q_2R_2$$

is the QR factorization of $T_m^{(1)}-\theta_2 I$. Postmultiplying by Q_2, we get

$$(A-\theta_2 I)(U_m^{(1)}Q_2) = (U_m^{(1)}Q_2)(R_2Q_2)+\beta_m u_{m+1}(b_{m+1}^{(1)H}Q_2).$$

It implies that

$$AU_m^{(2)} = U_m^{(2)}T_m^{(2)}+\beta_m u_{m+1}b_{m+1}^{(2)H},$$

where

$$U_m^{(2)} \equiv U_m^{(1)}Q_2$$

is orthonormal,

$$T_m^{(2)} \equiv R_2 Q_2 + \theta_2 I = \left[\begin{array}{c|cc} T_{m-2}^{(2)} & 0 & 0 \\ \hline & \theta_2 & 0 \\ & & \theta_1 \end{array}\right]$$

is tridiagonal with unreduced matrix $T_{m-2}^{(2)}$ and

$$\begin{aligned} b_{m+1}^{(2)H} &\equiv b_{m+1}^{(1)H} Q_2 = q_{m-1,m}^{(1)} e_{m-1}^H Q_2 + q_{m,m}^{(1)} e_m^\top Q_2 \\ &= \left[\begin{array}{cccccc} 0 & \cdots & 0 & \times & \times & \times \end{array}\right]. \end{aligned}$$

For on postmultiplying (16) by e_1, we get

$$(A - \theta_2 I) u_1^{(1)} = (A - \theta_2 I)(U_m^{(1)} e_1) = U_m^{(2)} R_2 e_1 = r_{11}^{(2)} u_1^{(2)}.$$

Since $H_m^{(1)}$ is unreduced, $r_{11}^{(2)}$ is nonzero. Therefore, the first column of $U_m^{(2)}$ is a multiple of

$$(A - \theta_2 I) u_1^{(1)} = 1/r_{11}^{(1)} (A - \theta_2 I)(A - \theta_1 I) u_1.$$

Repeating this process with $\theta_3, \ldots, \theta_{m-k}$, the result will be a Krylov decomposition

$$AU_m^{(m-k)} = U_m^{(m-k)} T_m^{(m-k)} + \beta_m u_{m+1} b_{m+1}^{(m-k)H}$$

with the following properties

1. $U_m^{(m-k)}$ is orthonormal.
2. $T_m^{(m-k)}$ is tridiagonal.
3. The first $k-1$ components of $b_{m+1}^{(m-k)H}$ are zero.
4. The first column of $U_m^{(m-k)}$ is a multiple of $(A - \theta_1 I) \cdots (A - \theta_{m-k} I) u_1$.

Corollary 24

Let $\theta_1, \ldots, \theta_m$ *be eigenvalues of* T_m*. If the implicitly restarted* QR *step is performed with shifts* $\theta_1, \ldots, \theta_{m-k}$*, then the matrix* $T_m^{(m-k)}$ *has the form*

$$T_m^{(m-k)} = \begin{bmatrix} T_{kk}^{(m-k)} & 0 \\ 0 & D^{(m-k)} \end{bmatrix},$$

where $D^{(m-k)}$ *is an digonal matrix with Ritz value* $\theta_1, \ldots, \theta_{m-k}$ *on its diagonal.*

For $k = 3$ and $m = 6$,

$$A\left[\begin{array}{ccc|ccc} u & u & u & u & u & u \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} u & u & u & u & u & u \end{array}\right] \left[\begin{array}{ccc|ccc} \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 & 0 \\ \hline 0 & 0 & \times & \times & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & \times \end{array}\right]$$

$$+u\left[\begin{array}{ccc|ccc} 0 & 0 & q & q & q & q \end{array}\right].$$

Therefore, the first k columns of the decomposition can be written in the form

$$AU_k^{(m-k)} = U_k^{(m-k)} T_{kk}^{(m-k)} + t_{k+1,k} u_{k+1}^{(m-k)} e_k^\top + \beta_m q_{mk} u_{m+1} e_k^\top,$$

where $U_k^{(m-k)}$ consists of the first k columns of $U_m^{(m-k)}$, $T_{kk}^{(m-k)}$ is the leading principal submatrix of order k of $T_m^{(m-k)}$, and q_{mk} is from the matrix $Q = Q_1 \cdots Q_{m-k}$.

Hence if we set

$$\begin{aligned}\tilde{U}_k &= U_k^{(m-k)},\\ \tilde{T}_k &= T_{kk}^{(m-k)},\\ \tilde{\beta}_k &= \|t_{k+1,k}u_{k+1}^{(m-k)} + \beta_m q_{mk} u_{m+1}\|_2,\\ \tilde{u}_{k+1} &= \tilde{\beta}_k^{-1}(t_{k+1,k}u_{k+1}^{(m-k)} + \beta_m q_{mk} u_{m+1}),\end{aligned}$$

then

$$A\tilde{U}_k = \tilde{U}_k\tilde{T}_k + \tilde{\beta}_k \tilde{u}_{k+1} e_k^\top$$

is a Lanczos decomposition whose starting vector is proportional to $(A - \theta_1 I) \cdots (A - \theta_{m-k} I)u_1$.

- Avoid any matrix-vector multiplications in forming the new starting vector.
- Get its Lanczos decomposition of order k for free.
- For large n the major cost will be in computing UQ.

Practical Implementation

Restarted Lanczos method

Input: Given Lanczos decomp. $AU_m = U_m T_m + \beta_m u_{m+1} e_m^\top$
Output: new Lanczos decomp. $AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top$

1: Compute the eigenvalues $\theta_1, \ldots, \theta_m$ of T_m.
2: Determine shifts, said $\theta_1, \ldots, \theta_{m-k}$, and set $b_m = e_m^\top$.
3: **for** $j = 1, \ldots, m-k$ **do**
4: Compute QR factorization: $T_m - \theta_j I = Q_m R_m$.
5: Update $T_m := R_m Q_m + \theta_j I$, $U_m := U_m Q_m$, $b_m := Q_m^\top b_m$.
6: **end for**
7: Compute $v = \beta_k u_{k+1} + \beta_m b_m(k) u_{m+1}$.
8: Set $U_k := U_m(:, 1:k)$, $\beta_k = \|v\|_2$, $u_{k+1} = v/\beta_k$, and $T_k := T_m(1:k, 1:k)$,

Question

Can we implicitly compute Q_m and get new tridiagonal matrix T_m?

General algorithm

1. Determine the first column c_1 of $T_m - \theta_j I$.
2. Let $\widehat{Q}$ be a Householder transformation such that $\widehat{Q}^\top c_1 = \sigma e_1$.
3. Set $T = \widehat{Q}^\top T_m \widehat{Q}$.
4. Use Householder transformation $\widetilde{Q}$ to reduce T to a new tridiagonal form $\widehat{T} \equiv \widetilde{Q}^\top T \widetilde{Q}$.
5. Set $Q_m = \widehat{Q}\widetilde{Q}$.

Question

General algorithm = one step of single shift QR algorithm ?

Answer:

(I) Let

$$T_m - \theta_j I = \begin{bmatrix} c_1 & C_* \end{bmatrix} = Q_m R_m = \begin{bmatrix} q & Q_{m*} \end{bmatrix} \begin{bmatrix} \rho & r_* \\ 0 & R_* \end{bmatrix}$$

be the QR factorization of $T_m - \theta_j I$. Then $c_1 = \rho q$. Partition $\widehat{Q} \equiv \begin{bmatrix} \hat{q} & \widehat{Q}_* \end{bmatrix}$, then $c_1 = \sigma \widehat{Q} e_1 = \sigma \hat{q}$ which implies that q and $\hat{q}$ are proportional to c_1.

(II) Since $\widehat{T} = \widetilde{Q}^\top T \widetilde{Q}$ is tridiagonal, we have

$$\widetilde{Q} e_1 = e_1.$$

Hence,

$$(\widehat{Q}\widetilde{Q}) e_1 = \widehat{Q} e_1 = \hat{q}$$

which implies that the first column of $\widehat{Q}\widetilde{Q}$ is proportional to q.

(III) Since $(\widehat{Q}\widetilde{Q})^\top T_m (\widehat{Q}\widetilde{Q})$ is tridiagonal and the first column of $\widehat{Q}\widetilde{Q}$ is proportional to q, by the implicit Q Theorem, if $\widehat{T}$ is unreduced, then $\widehat{Q} = Q_0 Q_1$ and $\widehat{T} = (\widehat{Q}\widetilde{Q})^\top T_m (\widehat{Q}\widetilde{Q})$.

Definition 25 (Givens rotation)

A plane rotation (also called a Givens rotation) is a matrix of the form

$$G = \begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix}$$

where $|c|^2 + |s|^2 = 1$.

Given $a \neq 0$ and b, set

$$v = \sqrt{|a|^2 + |b|^2}, \; c = |a|/v \text{ and } s = \frac{a}{|a|} \cdot \frac{\bar{b}}{v},$$

then

$$\begin{bmatrix} c & s \\ -\bar{s} & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} v\dfrac{a}{|a|} \\ 0 \end{bmatrix}.$$

Let

$$G_{ij} = \begin{pmatrix} I_{i-1} & & & & \\ & c & & s & \\ & & I_{j-i-1} & & \\ & -\bar{s} & & c & \\ & & & & I_{n-j} \end{pmatrix}.$$

(I) Compute the first column $t_1 \equiv \begin{bmatrix} \alpha_1 - \theta_j \\ \beta_1 \\ 0 \end{bmatrix}$ of $T_m - \theta_j I$ and determine Givens rotation G_{12} such that $G_{12}t_1 = \gamma e_1$.

(II) Set $T_m := G_{12} T_m G_{12}^\top$.

$$T_m := G_{12} \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix} G_{12}^\top = \begin{bmatrix} \times & \times & + & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ + & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

(III) Construct orthonormal Q such that $T_m := QT_mQ^\top$ is tridiagonal:

$$T_m := G_{23}\begin{bmatrix} \times & \times & + & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ + & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix} G_{23}^\top = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & + & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & + & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

$$T_m := G_{34}\begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & + & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & + & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix} G_{34}^\top = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & + \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & + & \times & \times \end{bmatrix}$$

$$T_m := G_{45}\begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & + \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & + & \times & \times \end{bmatrix} G_{45}^\top = \begin{bmatrix} \times & \times & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

Implicit Restarting for Lanczos method

Input: Given Lanczos decomp. $AU_m = U_m T_m + \beta_m u_{m+1} e_m^\top$
Output: new Lanczos decomp. $AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top$

1: Compute the eigenvalues $\theta_1, \ldots, \theta_m$ of T_m.
2: Determine shifts, said $\theta_1, \ldots, \theta_{m-k}$, and set $b_m = e_m^\top$.
3: **for** $j = 1, \ldots, m-k$ **do**
4: Compute Givens rotation G_{12} such that $G_{12} \begin{bmatrix} \alpha_1 - \theta_j \\ \beta_1 \\ 0 \end{bmatrix} = \gamma e_1$.
5: Update $T_m := G_{12} T_m G_{12}^\top$, $U_m := U_m G_{12}^\top$, $b_m := G_{12} b_m$.
6: Compute Givens rotations $G_{23}, \ldots, G_{m-1,m}$ such that $T_m := G_{m-1,m} \cdots G_{23} T_m G_{23}^\top \cdots G_{m-1,m}^\top$ is tridiagonal.
7: Update $U_m := U_m G_{23}^\top \cdots G_{m-1,m}^\top$ and $b_m := G_{m-1,m} \cdots G_{23} b_m$.
8: **end for**
9: Compute $v = \beta_k u_{k+1} + \beta_m b_m(k) u_{m+1}$.
10: Set $U_k := U_m(:, 1:k)$, $\beta_k = \|v\|_2$, $u_{k+1} = v/\beta_k$, and $T_k := T_m(1:k, 1:k)$,

Problem

- Mathematically, u_j must be orthogonal.
- In practice, they can lose orthogonality.

Solutions

Reorthogonalize the vectors at each step.

j-th step of Lanczos process

Input: Given β_{j-1} and orthonormal matrix $U_j = [u_1, \ldots, u_j]$.
Output: α_j, β_j and unit vector u_{j+1} with $u_{j+1}^\top U_j = 0$ and $Au_j = \alpha_j u_j + \beta_{j-1} u_{j-1} + \beta_j u_{j+1}$.

1: Compute $u_{j+1} = Au_j - \beta_{j-1} u_{j-1}$ and $\alpha_j = u_j^\top u_{j+1}$;
2: Update $u_{j+1} := u_{j+1} - \alpha_j u_j$;
3: **for** $i = 1, \ldots, j$ **do**
4: Compute $\gamma_i = u_i^\top u_{j+1}$ and update $u_{j+1} := u_{j+1} - \gamma_i u_i$;
5: **end for**
6: Update $\alpha_j := \alpha_j + \gamma_j$ and compute $\beta_j = \|u_{j+1}\|_2$ and $u_{j+1} := u_{j+1}/\beta_j$.

Lanczos algorithm with implicit restarting

Input: Given initial unit vector u_1, number k of desired eigenpairs, restarting number m and stopping tolerance ε.

Output: desired eigenpairs (θ_i, x_i) for $i = 1, \ldots, k$.

1: Compute Lanczos decomposition with order k:
   $AU_k = U_k T_k + \beta_k u_{k+1} e_k^\top$;
2: **repeat**
3: Extend the Lanczos decomposition from order k to order m:
   $AU_m = U_m T_m + \beta_m u_{m+1} e_m^\top$;
4: Use implicitly restarting scheme to reform a new Lanczos decomposition with order k;
5: Compute the eigenpairs (θ_i, s_i), $i = 1, \ldots, k$, of T_k;
6: **until** $(|\beta_k||s_{i,k}| < \varepsilon$ for $i = 1, \ldots, k)$
7: Compute eigenvector $x_i = U_k s_i$ for $i = 1, \ldots, k$.

Arnoldi method

Recall: Arnoldi decomposition of unsymmetric A:

$$AU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^\top, \tag{17}$$

where H_k is unreduced upper Hessenberg.
Write (17) in the form

$$Au_k = U_k h_k + h_{k+1,k} u_{k+1}.$$

Then from the orthogonality of U_{k+1}, we have

$$h_k = U_k^H A u_k.$$

Since $h_{k+1,k} u_{k+1} = Au_k - U_k h_k$ and $\|u_{k+1}\|_2 = 1$, we must have

$$h_{k+1,k} = \|Au_k - U_k h_k\|_2, \quad u_{k+1} = h_{k+1,k}^{-1}(Au_k - U_k h_k).$$

Arnoldi process

1: **for** $k = 1, 2, \ldots$ **do**
2: $\quad h_k = U_k^H A u_k$;
3: $\quad v = A u_k - U_k h_k$;
4: $\quad h_{k+1,k} = \|v\|_2$;
5: $\quad u_{k+1} = v / h_{k+1,k}$;
6: $\quad \hat{H}_k = \begin{bmatrix} \hat{H}_{k-1} & h_k \\ 0 & h_{k+1,k} \end{bmatrix}$;
7: **end for**

- The computation of u_{k+1} is actually a form of the well-known Gram-Schmidt algorithm.
- In the presence of inexact arithmetic cancelation in statement 3 can cause it to fail to produce orthogonal vectors.
- The cure is process called reorthogonalization.

Reorthogonalized Arnoldi process

for $k = 1, 2, \ldots$ **do**

$h_k = U_k^H A u_k$;

$v = A u_k - U_k h_k$;

$w = U_k^H v$;

$h_k = h_k + w$;

$v = v - U_k w$;

$h_{k+1,k} = \|v\|_2$;

$u_{k+1} = v / h_{k+1,k}$;

$\hat{H}_k = \begin{bmatrix} \hat{H}_{k-1} & h_k \\ 0 & h_{k+1,k} \end{bmatrix}$;

end for

Let $y_i^{(k)}$ be an eigenvector of H_k associated with the eigenvalue $\lambda_i^{(k)}$ and $x_i^{(k)} = U_k y_i^{(k)}$ the Ritz approximate eigenvector.

Theorem 26

$$(A - \lambda_i^{(k)} I) x_i^{(k)} = h_{k+1,k} e_k^\top y_i^{(k)} u_{k+1}.$$

and therefore,

$$\|(A - \lambda_i^{(k)} I) x_i^{(k)}\|_2 = |h_{k+1,k}| \, | e_k^\top y_i^{(k)} |.$$

Generalized eigenvalue problem

Consider the generalized eigenvalue problem

$$Ax = \lambda Bx,$$

where B is nonsingular. Let

$$C = B^{-1}A.$$

Applying Arnoldi process to matrix C, we get

$$CU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^\top,$$

or

$$AU_k = BU_k H_k + h_{k+1,k} B u_{k+1} e_k^\top. \tag{18}$$

Write the k-th column of (18) in the form

$$Au_k = BU_k h_k + h_{k+1,k} Bu_{k+1}. \tag{19}$$

Let U_k satisfy that

$$U_k^\top U_k = I_k.$$

Then

$$h_k = U_k^\top B^{-1} Au_k$$

and

$$h_{k+1,k} u_{k+1} = B^{-1} Au_k - U_k h_k \equiv t_k$$

which implies that

$$h_{k+1,k} = \|t_k\|, \quad u_{k+1} = h_{k+1,k}^{-1} t_k.$$

Shift-and-invert Lanczos for GEP

Consider the generalized eigenvalue problem

$$Ax = \lambda Bx, \tag{20}$$

where A is symmetric and B is symmetric positive definite.

Shift-and-invert

- Compute the eigenvalues which are closest to a given shift value σ.
- Transform (20) into

$$(A - \sigma B)^{-1} Bx = (\lambda - \sigma)^{-1} x. \tag{21}$$

The basic recursion for applying Lanczos method to (21) is

$$(A - \sigma B)^{-1} B U_j = U_j T_j + \beta_j u_{j+1} e_j^\top, \tag{22}$$

where the basis U_j is B-orthogonal and T_j is a real symmetric tridiagonal matrix defined by

$$T_j = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{j-1} \\ & & \beta_{j-1} & \alpha_j \end{bmatrix}$$

or equivalent to

$$(A - \sigma B)^{-1} B u_j = \alpha_j u_j + \beta_{j-1} u_{j-1} + \beta_j u_{j+1}.$$

By the condition $U_j^* B U_j = I_j$, it holds that

$$\alpha_j = u_j^* B(A - \sigma B)^{-1} B u_j, \quad \beta_j^2 = t_j^* B t_j,$$

where

$$t_j \equiv (A - \sigma B)^{-1} B u_j - \alpha_j u_j - \beta_{j-1} u_{j-1} = \beta_j u_{j+1}.$$

An eigenpair (θ_i, s_i) of T_j is used to get an approximate eigenpair (λ_i, x_i) of (A, B) by

$$\lambda_i = \sigma + \frac{1}{\theta_i}, \quad x_i = U_j s_i.$$

The corresponding residual is

$$\begin{aligned} r_i &= AU_j s_i - \lambda_i B U_j s_i = (A - \sigma B) U_j s_i - \theta_i^{-1} B U_j s_i \\ &= -\theta_i^{-1} \left[B U_j - (A - \sigma B) U_j T_j \right] s_i \\ &= -\theta_i^{-1} (A - \sigma B) \left[(A - \sigma B)^{-1} B U_j - U_j T_j \right] s_i \\ &= -\theta_i^{-1} \beta_j (e_j^\top s_i)(A - \sigma B) u_{j+1} \end{aligned}$$

which implies $\|r_i\|$ is small whenever $|\beta_j (e_j^\top s_i)/\theta_i|$ is small.

Shift-and-invert Lanczos method for symmetric GEP

Given starting vector t, compute $q = Bt$ and $\beta_0 = \sqrt{|q^* t|}$.
for $j = 1, 2, \ldots,$ **do**
  Compute $w_j = q/\beta_{j-1}$ and $u_j = t/\beta_{j-1}$.
  Solve linear system $(A - \sigma B)t = w_j$.
  Set $t := t - \beta_{j-1} u_{j-1}$; compute $\alpha_j = w_j^* t$ and reset $t := t - \alpha_j u_j$.
  B-reorthogonalize t to $u_1, \ldots, u_j$ if necessary.
  Compute $q = Bt$ and $\beta_j = \sqrt{|q^* t|}$.
  Compute approximate eigenvalues $T_j = S_j \Theta_j S_j^*$.
  Test for convergence.
end for
Compute approximate eigenvectors $X = U_j S_j$.

Krylov-Schur restarting method (Locking and Purging)

Theorem 27 (Schur decomposition)

Let $A \in \mathbb{C}^{n\times n}$. Then there is a unitary matrix V such that

$$A = VTV^H,$$

where T is upper triangular.

The process of Krylov-Schur restarting:

- Compute the Schur decomposition of the Rayleigh quotient
- Move the desired eigenvalues to the beginning
- Throw away the rest of the decomposition

Use Lanczos process to generate the Lanczos decomposition of order $m \equiv k + p$

$$(A - \sigma B)^{-1} B U_{k+p} = U_{k+p} T_{k+p} + \beta_{k+p} u_{k+p+1} e_{k+p}^{\top}. \tag{23}$$

Let

$$T_{k+p} = V_{k+p} D_{k+p} V_{k+p}^{\top} \equiv \begin{bmatrix} V_k & V_p \end{bmatrix} \text{diag}\,(D_k, D_p) \begin{bmatrix} V_k^{\top} \\ V_p^{\top} \end{bmatrix} \tag{24}$$

be a Schur decomposition of T_{k+p}. The diagonal elements of D_k and D_p contain the k wanted and p unwanted Ritz values, respectively.

Substituting (24) into (23), it holds that

$$\begin{aligned} & (A-\sigma B)^{-1}B(U_{k+p}V_{k+p}) \\ = \quad & (U_{k+p}V_{k+p})(V_{k+p}^{\top}T_{k+p}V_{k+p}) + \beta_{k+p}u_{k+p+1}(e_{k+p}^{\top}V_{k+p}), \end{aligned}$$

which implies that

$$(A-\sigma B)^{-1}B\tilde{U}_k = \tilde{U}_k D_k + u_{k+p+1}t_k^{\top} = \begin{bmatrix} \tilde{U}_k & \tilde{u}_{k+1} \end{bmatrix} \begin{bmatrix} D_k \\ t_k^{\top} \end{bmatrix}$$

is a Krylov decomposition of order k where $\tilde{U}_k \equiv U_{k+p}V_k$, $\tilde{u}_{k+1} = u_{k+p+1}$ and $[t_k^{\top}, t_p^{\top}] \equiv \beta_{k+p}e_{k+p}^{\top}V_{k+p}$. The new vectors $\tilde{u}_{k+2}, \ldots, \tilde{u}_{k+p+1}$ are computed sequentially starting from $\tilde{u}_{j+2}$ with

$$(A-\sigma B)^{-1}B\tilde{U}_{k+1} = \begin{bmatrix} \tilde{U}_{k+1} & \tilde{u}_{k+2} \end{bmatrix} \begin{bmatrix} D_k & t_k \\ t_k^{\top} & \tilde{\alpha}_{k+1} \\ 0 & \tilde{\beta}_{k+1} \end{bmatrix},$$

or equivalently

$$(A - \sigma B)^{-1} B\tilde{u}_{k+1} = \tilde{U}_k t_k + \tilde{\alpha}_{k+1}\tilde{u}_{k+1} + \tilde{\beta}_{k+1}\tilde{u}_{k+2},$$

for

$$\tilde{\alpha}_{k+1} = \tilde{u}_{k+1}^* B(A - \sigma B)^{-1} B\tilde{u}_{k+1}.$$

Consequently a new Krylov decomposition of order $k + p$ can be generated by

$$(A - \sigma B)^{-1} B\tilde{U}_{k+p} = \tilde{U}_{k+p}\tilde{T}_{k+p} + \tilde{\beta}_{k+p}\tilde{u}_{k+p+1} e_{k+p}^{\top}, \tag{25}$$

where

$$\tilde{T}_{k+p} = \begin{bmatrix} D_k & t_k & & & \\ t_k^{\top} & \tilde{\alpha}_{k+1} & \tilde{\beta}_{k+1} & & \\ & \tilde{\beta}_{k+1} & \tilde{\alpha}_{k+2} & \ddots & \\ & & \ddots & \ddots & \tilde{\beta}_{k+p-1} \\ & & & \tilde{\beta}_{k+p-1} & \tilde{\alpha}_{k+p} \end{bmatrix}. \tag{26}$$

Shift-and-invert Krylov-Schur method for solving $Ax = \lambda Bx$

Build an initial Lanczos decomposition of order $k + p$:

$$(A - \sigma B)^{-1} BU_{k+p} = U_{k+p}T_{k+p} + \beta_{k+p}u_{k+p+1}e_{k+p}^{\top}.$$

repeat

Compute Schur decomposition $T_{k+p} = \begin{bmatrix} V_k & V_p \end{bmatrix} \text{diag}\left(D_k, D_p\right) \begin{bmatrix} V_k & V_p \end{bmatrix}^{\top}$.

Set $U_k := U_{k+p}V_k$, $u_{k+1} := u_{k+p+1}$ and $t_k^{\top} := \beta_{k+p}e_{k+p}^{\top}V_k$ to get

$$(A - \sigma B)^{-1} BU_k = U_k D_k + u_{k+1}t_k^{\top}.$$

for $j = k + 1, \ldots, k + p$ **do**

Solve linear system $(A - \sigma B)q = Bu_j$.

if $j = k + 1$ **then**

Set $q := q - U_k t_k$.

else

Set $q := q - \beta_{j-1}u_{j-1}$.

end if

Compute $\alpha_j = u_j^* Bq$ and reset $q := q - \alpha_j u_j$.

B-reorthogonalize q to $u_1, \ldots, u_{j-1}$ if necessary.

Compute $\beta_j = \sqrt{q^* Bq}$ and $u_{j+1} = q/\beta_j$.

end for

Decompose $T_{k+p} = S_{k+p}\Theta_{k+p}S_{k+p}^*$ to get the approximate eigenvalues.

Test for convergence.

until all the k wanted eigenpairs are convergent